Fig. 1

```
                  10                    30                    50
GTGAAGAACGAAAAAACCTTCTTTGAAGAGCTTTACGAGGCTTTAGAGGAAACCCACGAC
 M  K  N  E  K  T  F  F  E  E  L  Y  E  A  L  E  E  T  H  D
                  70                    90                   110
AACACCGATGCCACTAGGGGGTCAGATAGGGGGTCAGAGGACTTCTTCTTGGCCACCGAC
 N  T  D  A  T  R  G  S  D  R  G  S  E  D  F  F  L  A  T  D
                 130                   150                   170
CCCCCTCCAGATGGAGGTGCCGAAAATCGCCTCGCGAAGGGCTTTACATACCAAAAAGAG
 P  P  P  D  G  G  A  E  N  R  L  A  K  G  F  T  Y  Q  K  E
                 190                   210                   230
GCACTTAGGATTGCTTTACCCGAGAAAGACCATGAGGCTTTCCTTTCCTCTGTTGGGGCC
 A  L  R  I  A  L  P  E  K  D  H  E  A  F  L  S  S  V  G  A
                 250                   270                   290
CCCCCTATACCACCAGCTGAACCCCCCGTTGGGAATGTATGTCAAGCCGTCCAGGACGGG
 P  P  I  P  P  A  E  P  P  V  G  N  V  C  Q  A  V  Q  D  G
                 310                   330                   350
CCTCAGAAGCTTCTGGAACTCCTCCAGGAGATTGCCCGCTCCACCATCCCCTACGGCAAC
 P  Q  K  L  L  E  L  L  Q  E  I  A  R  S  T  I  P  Y  G  N
                 370                   390                   410
CGGGAGCTCTGGAGGAAGGTGGGGACGGTCGTCTTCATGGTCCCCCTGGAGATGTTGGCC
 R  E  L  W  R  K  V  G  T  V  V  F  M  V  P  L  E  M  L  A
                 430                   450                   470
CTCAACCTGGGGGTCACCCGGCAGACCGTCCACGCCTGGAAGAAGGTCCTTGAGAAAAAG
 L  N  L  G  V  T  R  Q  T  V  H  A  W  K  K  V  L  E  K  K
                 490                   510                   530
GGCCTGGTGGCCACCGACGTCCTTCACCAAACCGTCAACGGGGAGCGCCGGGCCATCGGC
 G  L  V  A  T  D  V  L  H  Q  T  V  N  G  E  R  R  A  I  G
                 550                   570                   590
ACCCTTTGGGCCGTCCGGCTGAGGCCAGGGAAAGCCAGGCTCACCCTGGACGACTACATC
 T  L  W  A  V  R  L  R  P  G  K  A  R  L  T  L  D  D  Y  I
                 610                   630                   650
TACCCCTGGAGGAACCTCGCCCTAGACATGGCCAACGGCGTGCTCTCCTTCAACTGGGTC
 Y  P  W  R  N  L  A  L  D  M  A  N  G  V  L  S  F  N  W  V
                 670                   690                   710
AAGGCCTACCAGGACCACGGAATCCGCCCCACCCTGGACGTGCTGGTCCTCTGGGCTCAG
 K  A  Y  Q  D  H  G  I  R  P  T  L  D  V  L  V  L  W  A  Q
                 730                   750                   770
GGGAAAAGGGTGATGCCCAACACCAAGACCGTGGCCGTTGACCTGGGCCTCATCCTGGTC
 G  K  R  V  M  P  N  T  K  T  V  A  V  D  L  G  L  I  L  V
                 790                   810                   830
CTCCCCGAGGTGGAGCGTTCCAAACTCCCGGCCCTTATCACCCTCATTGCTACGTACATT
 L  P  E  V  E  R  S  K  L  P  A  L  I  T  L  I  A  T  Y  I
                 850                   870                   890
GCCGATCTCCTAGATGACCGTCGTTCAAGACGTTTCTATGCAGGCTTGCTGTGGGCTGTG
 A  D  L  L  D  D  R  R  S  R  R  F  Y  A  G  L  L  W  A  V
                 910                   930                   950
GCCAGGGGTGAACTCCCCGCGCAATATCTATTTGCCGTCCTAATGCGGGTTATCCGAGAT
 A  R  G  E  L  P  A  Q  Y  L  F  A  V  L  M  R  V  I  R  D
                 970                   990                  1010
TACACGGATGGCCATCTGACACGACCGGGAGCGTACCTAGTGAAGACCCTCAAGGAGGCC
 Y  T  D  G  H  L  T  R  P  G  A  Y  L  V  K  T  L  K  E  A

TCCTGA
 S  *
```

Fig. 2

```
  1   CTATAACGGCCTTTTAGGAGGGGGGGATTGCCAGCCGCTGGGCTGACGGTTATTTTGGACC

 61   CATAAAAAGGCGAAACCGAGGCGGTTGCCCCGGATCACCCCCAAGACCTAGGGTAACGCC

121   TCGGGCTCCAGATGACAAGGAGGTCCGAGGGTGAAGAACGAAAAAACCTTCTTTGAAGAG
                            M   K   N   E   K   T   F   F...(RepT)
```

Fig. 3

```
   1 tctagaaggt cagggtggac aaggaaaaca ccatagcccc tgccaagaag atggacgagt
  61 tggtgtccgg aaaagtggcc atccggggcg ctcttgacaa ctatttttcca gcggtggcca
 121 ccggcattgg ccacgaggta cgagcttgtg gagtagacgg ccacaaaggg gtcgtcctca
 181 aacttctttt ctagtgccgc ttggacgaag gggaggaaga ggaaaggctt catggcctca
 241 cctccttccc ctcctccttg gcggccttag cggcgtaaaa ctctgagacg gcctgaagtt
 301 tagggatttc gctttcgggg ataagaatcc ggcggctcag gggatgccgg atggccctta
 361 tcctgccgtc ccttatgtac tcgtaaatgg tggccttggg tactttaaac cgttctgaaa
 421 cttctctaac agagagcaca aaacctctaa aaacctatca atcccaccga ttccagtata
 481 ccataaatgg cacaaagttt tgagaaggtg gtcaaacaaa aaggctttct cggtcaggtt
 541 atggtgaggt gggggcggtc aaaggccgac ttaagtttgg taaagccggg aggaagcaaa
 601 ccggggtgtt accatgcaac agatggccga gtggaacgtg tggacacaga gaagcgttga
 661 gcttctggag aaggggtatt tggataaact actgcaggtc tataaagggg aaagtggctc
 721 ttcgaggtca gtaccagagg aggtagagga aaaacttcgc gaggcctaca aggcatacga
 781 ggggaggcag gatagtccgg aggcagaaac gaaactcgtg gaagccgtgc taaatgccag
 841 aaaaaaggtc gagcggtccc ccttcaatca cccctacctg cctttggtct actacctggt
 901 ttcggaaaaa gcagaaaaag cgaacaaggc ccttgaggag gcattgcagg aggttgcctc
 961 aaagcaccca gaaaccatcc gcgtcctggc caaggaagcg caaagaagag gcgtagaagc
1021 cttgatccaa aggctcaagg agcctcccga aataaatcgg cagatagggc cgatgttcaa
1081 aaggtggtac aaagaagagc taaaggggaa aatagaagag aggcttccag gccctaccaa
1141 accaaagatt gtggtagtat cccctgaaaa aagtaaaccg gagcaagcac cccttattgc
1201 ggagagagaa gcgggcatca tcatatacac gggatcggat gaagctttga aagatgccgc
1261 caaggaaaac ctgggccttg gcgaggaagc agaactaggc accaagggcg tagatttcta
1321 cgtggtcatc cggcgtagcc ctgaagagac atggcaccta acaggagaag tgaagtttca
1381 atccgacttt ggcggaaacc aagacaacca gaaactagta gcaaaggctt ccataaggtt
1441 ggaccttgag aagaggcaca taggaatagt ggtggtggac ggaatgcctg tggtgagcaa
1501 gtttcgtggg tgggccggac tggggaaaga aacgatcgtt acatccgtac tcctccttcc
1561 agacctgata gcggagctct accaaaaggg tgaagaagcc ctgggcctct agaaggcgga
1621 cacaatctca aacttgtgct gtagcctggg gaaatcctct aacacccttc tagtgaaggc
1681 tttgaccgcc tcccaggagg catctatgcc gatggatcgc cgctttaaga ggggtgaggc
1741 tataagcgta gtaccggagc ctgcgaaggg atcgagcact aaatcccccct cgttactccc
1801 tgtttggacg atgagcttga gcatgtccag attttctcg gtggggtatc gcgggtacgg
1861 aggatccttg aactgccaaa cgtcctggag cttcttcccc ttcttcaggc gatcccgagc
1921 gtaaactttc ttccgcggca ccccgttctt tgaccagaca ataagccctt gagcgtctag
1981 ctcgtcaagc ttctccgggg gatagcgcca atgccgtcca ggaggggaa gtattcctcg
2041 ccaaggcctt ccggtagggc catccttggt ttctccagga gcatgcaggg gattggtggt
2101 gtaccgttcc ccgttctcgt ctacaaaggg gaaaagccta gcgatctcct cttccgaata
2161 ggggctagcc gattcgttcc aaacgtagtc ccgcgttttg gagtagacga ggatcatgtc
2221 cttttgcgat ccgaaggcct tacgggaaaa gttttttggga tttgaagcga tgcgggcgat
2281 atggttaacg aagtttcgcc ggccaaagac ctcatcaagg atgagcttca cctcgaaccc
2341 gtatttctcg tctatgtgaa cgaagatcag tcctgagtcc gccatcagct ccctgagaag
2401 tatcaagcgc tccctcagga actccacaaa ctgaggacca tcgagggtgt catcgtagcc
2461 caactgaccg ttttgggct ggctgacggt agcaacgcga tctgtttcat cgccgccaac
2521 gagaaactgc tggccggttc cataaggcgg gtcaatatag accaactgga ccttccccgc
2581 atacccacca ggctcccgga gcatccaccg gagaacctga ccgttttccc ccaaaaagta
2641 ggtgccaata ggatcaatct caaaaagggg ggcatttccc cctaggaaga ggagggtttc
2701 ttttcgcaaa acaagttgtg gggtgggctg atcaagaatc tccttctcat cgcgttttcc
2761 ggggtagacc aacctaaagg gcgaaggttc cgaggttttc gaggctttca aggggggcttt
2821 tcgggtcaaa ccagggtagc tacggctcat tcttccctcc ccacagcgct cttaagcagg
2881 acctcatcac ccacaaccct cacgcactcc aaccaaggaa tccgccaaag gcggcctacc
2941 ttttgagccc gtatcttccc ctgacgtata gaccttcgga tcgtctcagg gtgcacccga
3001 aggatgtctg caagctcctc gggggtcagg tacacgggct tcatcctcat gacacaacct
3061 tacccacag aggacaacac atgcaactat gggcaaagta gacaacgaga ccaaaagctt
3121 gggccactct ctcaggaggc ctccttgagg gtcttcacta ggtacgctcc cggtcgtgtc
3181 agatggccat ccgtgtaatc tcggataacc cgcattagga cggcaaatag atattgcgcg
3241 gggagttcac ccctggccac agcccacagc aagcctgcat agaaacgtct tgaacgacgg
3301 tcatctagga gatcggcaat gtacgtagca atgagggtga taagggccgg gagtttggaa
```

Fig. 3 (continued)

```
3361 cgctccacct cggggaggac caggatgagg cccaggtcaa cggccacggt cttggtgttg
3421 ggcatcaccc tttttcccctg agcccagagg accagcacgt ccagggtggg gcggattccg
3481 tggtcctggt aggccttgac ccagttgaag gagagcacgc cgttggccat gtctagggcg
3541 aggttcctcc aggggtagat gtagtcgtcc agggtgagcc tggctttccc tggcctcagc
3601 cggacggccc aaagggtgcc gatggcccgg cgctccccgt tgacggtttg gtgaaggacg
3661 tcggtggcca ccaggccctt tttctcaagg accttcttcc aggcgtggac ggtctgccgg
3721 gtgacccca ggttgagggc caacatctcc aggggacca tgaagacgac cgtccccacc
3781 ttcctccaga gctcccggtt gccgtagggg atggtggagc gggcaatctc ctggaggagt
3841 tccagaagct tctgaggccc gtcctggacg gcttgacata cattcccaac gggggggttca
3901 gctggtggta taggggggggc cccaacagag gaaaggaaag cctcatggtc tttctcgggt
3961 aaagcaatcc taagtgcctc tttttggtat gtaaagccct tcgcgaggcg attttcggca
4021 cctccatctg gaggggggtc ggtggccaag aagaagtcct ctgaccccct atctgacccc
4081 ctagtggcat cggtgttgtc gtgggtttcc tctaaagcct cgtaaagctc ttcaaagaag
4141 gttttttcgt tcttcaccct cggacctcct tgtcatctgg agcccgaggc gttaccctag
4201 gtcttggggg tgatccggggg caaccgcctc ggtttcgcct ttttatgggt ccaaaataac
4261 cgtcagccca gcggctggca atcccccctc ctaaaaggcc gttataggcc ctgctaggag
4321 gggggtagta cttttcctacc cccctaggct tggagaggcc ttaggaggtc tcctagggcc
4381 tcgtggggggt gtaggggtaa cctcatggcc aggccggccg gctcgggact ctggaggagg
4441 cctccatagc ctactcgtgg tggaggtttg tgaagggggtt cactaatgca tacggctagc
4501 ctcgggatca cggccaaatg gtatgcaggt tttggtataa aaccctcagg tttgaggcta
4561 gtttatgtcg gttttatgca cctttgactc ggatcacggg cataaacacc agtttcctgc
4621 acgaaagaaa actttcgcga tctaagaggg ggaaagaggt gtagagggac ggccttcatg
4681 aaagttggcc tcttaggagg ccgttgtaga gggccgtctc gggttcaaat cctttccctc
4741 tctctccagg tttccgaggt tcgaggtctt ggtccaggtc ttgtaccaag ttttttgacca
4801 aagtctattc tcggaatata ggggtatctt gtctatcttc cctacgggat atctctgtct
4861 gtgtgaactt gatcccatcc caatacatat ctcaatctcc taatctcctc ttctctccag
4921 atccctaatc tcttcttcta cctctttctc ctcccaatta agaatggaga ggaaaaaccc
4981 cgaccagaac gagcttctcg gggtcagttt cggtaatctc gggacaggtt ttcatcgtct
5041 aggacgagga ttagggcatg aaaaatgggc tttgacaaaa tctttctaaa aaatactccc
5101 cgaggttggg gaagtgccct cggggagaag atttttggca gtttagatgt tatgctctat
5161 cacgggccgg aggcctccac gataagttgt cttggccaag taccgggcca ggtcgggggt
5221 gctcttcagc gtggtgatgg tactttcacg gaagttcaca agtccttta gaggcttcag
5281 gtcgggggata gtgctcaagt actcccaagc gttctcgggc ccgtggtcgg ggagaaggac
5341 aaaggggtcg ggcaaaagtt catctttgta cttaggacgg attactttag cacctgataa
5401 cttcagggcc gttaagaagg gcctcacctc ggagacgggt ggaaggagga cgtgggcgtg
5461 gaagaagacg aaccccgatt tttgggaagt ctccctccag tttgatgatg aacgttggga
5521 ggaagccggc caggatgtct ttcatcgcgc ctcgaacctc ggacacataa aaaactttcg
5581 tgtttgtcag ggcaagagtg ctatgtatga ggtaaccttc gggagtacaa agtgcctcaa
5641 gccgcctttc ccaacgctcc aaaactctag ggtcaggtgg tttaggtttt ctgaaaaact
5701 ctagcttttc agtggtcatt cctcacccct ctagcacgta ctctggaagg taaacctttg
5761 acacagcggc caagtctagc gtctcccagt ccagttggtc tgggacgcgt gagaagggga
5821 ggggcttggt gtagaggacc agaagaccc
```

pTsp45s
(5.8 kb)

XbaI
HindIII
SphI
NheI
HindIII
SphI
NheI
tsp45IM
HindIII
repT
HindIII
ScaI
ScaI
NheI
HindIII
tsp45IR
PstI
XbaI
ScaI
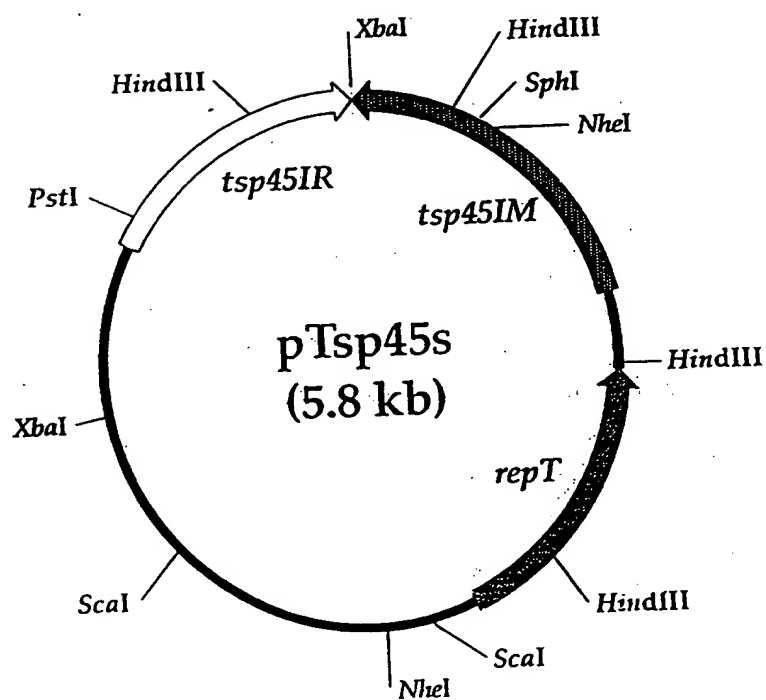
Fig. 4

Fig. 5

```
  1 ATGATCGTGGCTGTCACCGGCTTCAAGGGAGGGGTGGGGAAGACCACCACGGCGGTCCAC
    M  I  V  A  V  T  G  F  K  G  G  V  G  K  T  T  T  A  V  H

 61 CTGGCCTGCTTCCTGGCCGAGCGGGGCCCCACCCTGCTGGTGGACGGGGACCCCAACCGC
    L  A  C  F  L  A  E  R  G  P  T  L  L  V  D  G  D  P  N  R

121 TCCGCCACGGGGTGGCACCGGAGGGGAGGCCTCCCGGTGACCGTGGTGGACGAGCGGGTG
    S  A  T  G  W  H  R  R  G  G  L  P  V  T  V  V  D  E  R  V

181 GCGGCCCGGTACGCCCGGGAGCACGCCCACGTGGTCATAGACACCCAGGCCCGCCCCACG
    A  A  R  Y  A  R  E  H  A  H  V  V  I  D  T  Q  A  R  P  T

241 GAAGAGGACCTCCGGGCCCTCGCCAAGGGGGTGGACCTGCTGGTCCTGCCCACGTCCCCC
    E  E  D  L  R  A  L  A  K  G  V  D  L  L  V  L  P  T  S  P

301 GACGCCCTGGCCCTGGAGGCCCTCCTGGCCACCCTGGAAGCCCTGCGGGGGGCGGAGGCC
    D  A  L  A  L  E  A  L  L  A  T  L  E  A  L  R  G  A  E  A

361 CGCTTCCGGGTCCTCCTGACCATGGTGCCCCCGCCCCCGAGCCGGGACGGGGAGGAGGCC
    R  F  R  V  L  L  T  M  V  P  P  P  P  S  R  D  G  E  E  A

421 CGGGCCCTCTTGGGGGCGGAGGGCGTTCCCCTCTTCACAGGCTGGGTGAGGCGGGCGGCA
    R  A  L  L  G  A  E  G  V  P  L  F  T  G  W  V  R  R  A  A

481 GCCTTCCCCAAGGCCGCCCTCCTGGGGGTGCCTGTCTACCGGGTGCCCGACCCCAGGGCG
    A  F  P  K  A  A  L  L  G  V  P  V  Y  R  V  P  D  P  R  A

541 AGGCTGGCCTGGGGGGACTACGCGCGGGTGGGGGAAGAGCTCCTGAAGGAGGTGGGGGGA
    R  L  A  W  G  D  Y  A  R  V  G  E  E  L  L  K  E  V  G  G

601 TGA 603
```

a.

5' b.

c.

d.

3' e.

f

ORF 1
ORF 2
ORF 3
ORF 4
ORF 5
ORF 6
ORF 7

5,000

10,000

3'

5

Fig. 6

Fig. 7

```
         CTTATACACACAAACTATACACGTCTCTATCGGGCTTTTCTTAGCGCCATGTAAAACACC
   1     ---------+---------+---------+---------+---------+---------+    60
         CCTCCCATCTCCGGGTGTTTACAGCGGATACGGGAGGTTCAGCGGGAACTTTTCCCCTTG
  61     ---------+---------+---------+---------+---------+---------+   120
         TTGAAACTTTGGGGTCTGAGGCTCAACAGCAGAACAGCTTAGGTTGACTCAACACAGCTC
 121     ---------+---------+---------+---------+---------+---------+   180
         ATAAGTCCCTTCATTATCGCCTGAGTCAACCTATGAGTTAACCTTTTTTCAAGAAAAAGA
 181     ---------+---------+---------+---------+---------+---------+   240
         GATAAGTGAGTTTTGTCCTCTAGCACGACTTTTTTCTTTGAGTCAACCTCTGTGCCGACC
 241     ---------+---------+---------+---------+---------+---------+   300
         CCCCCGATTTTGAGTCAACCCCCCTTTGAGCCGAAACTTTGTTGGCACAGGGGTTGACTC
 301     ---------+---------+---------+---------+---------+---------+   360
         AGGGGTTGACTCAACGCGAATGGCCTCTGGAAGGGCGTTGAGCCGACCCCTCCCTCGTGT
 361     ---------+---------+---------+---------+---------+---------+   420
         GCCGACCCCCGCTCCACTATGAGCAGGGGGGAAAGTTACGGGAAAAGTTCCCCAAGTCCC
 421     ---------+---------+---------+---------+---------+---------+   480
         CCTTGACAAAAGATGACAATCGAGTTAATGTCACAGCGATGCGTCACTCACCTCTGGCTG
 481     ---------+---------+---------+---------+---------+---------+   540
         GGCTCACCCAGATGCGTGCGCGAACGTTTCAGAGCCTCCTTCGATTCCTGGCCAGGGAGG
 541     ---------+---------+---------+---------+---------+---------+   600
         GGCGCTACCCCACTGGTGTAGAGCTCGCCAAGGTGCTGGGGCGCAGCCCGCACGCCACGT
 601     ---------+---------+---------+---------+---------+---------+   660
         GGGCCATGCTCAGGGCTTTGACCCGTCATGGACTCGTGGAACGGCACGAGGGGGTCTATG
 661     ---------+---------+---------+---------+---------+---------+   720
         TTCTGACCCCTGCGGGCGTAGAACTTGCCAGGACCCTGGGAACCACCGTGTGGCGTGGGG
 721     ---------+---------+---------+---------+---------+---------+   780
         ATGAGGAGGTACAGACGGCGTTACAGCTGCTAGGAGTCGGTCATGCCGCCGAGGACAGGC
 781     ---------+---------+---------+---------+---------+---------+   840
         GCTGAAGCTTTTGAGCCGGGGCCCTCACCCAAGGCCACCCGGCTCCTCTCCCCTGGGAT
 841     ---------+---------+---------+---------+---------+---------+   900
         CCCAAATGGATCCCTCAGCGCCATTATCCTCCTGGCGGTCCTATAGCGCAAGGAGGTAGT
 901     ---------+---------+---------+---------+---------+---------+   960
         GGTGACGAAACACACAAATGTTTCACCCCACCTTTTGGATGCCGTAGAGGAGCTCGCTCG
 961     ---------+---------+---------+---------+---------+---------+  1020
         CCAGATTGCTGAAACCGCTAACAAGGCTTATTCCAGCCATTTCAGGCAGATTGTCAAAGT
1021     ---------+---------+---------+---------+---------+---------+  1080
         CCTGCCGCCTGAGGTTCCCGACCTCTACGCCTGGCTGGCCGCCCTGGATGACTCCGCCAT
1081     ---------+---------+---------+---------+---------+---------+  1140
         CGAGGAGCTTGCCCAGCGCCTGAGGGAGGTCGAGGGAAGCCCCCGCCCCCATTTCACCGC
1141     ---------+---------+---------+---------+---------+---------+  1200
         CGCCCTCAAAAAGGCCCTGGCCATCGCCCTACAGCGGCGGACCCTCGCCGAGATGCCCCC
1201     ---------+---------+---------+---------+---------+---------+  1260
         CACGTTCGCCAACGCGCTCCGCTGGGCGATGGAACGGCAAGGGGTGAGCATCCGCAAGCT
1261     ---------+---------+---------+---------+---------+---------+  1320
         TGCGAGAGAGGTAGGGGTCAGCAAAACCACTGTTAAAAAGTGGCGTGGAGGCCGCTTTGT
1321     ---------+---------+---------+---------+---------+---------+  1380
         CCCTCGTTCACGGACCTACGTGAGGAGGTTGGAGGAGATCCTGGACCTCCCGGAAGGCGC
1381     ---------+---------+---------+---------+---------+---------+  1440
         CCTTTCGGGACGACTACCCCGCTGGGGGTTGCCAAAAATATTGGAAGGTGTTGAGGGGAA
1441     ---------+---------+---------+---------+---------+---------+  1500
         AGATGCCCCTTATCCCGGGTTCACGCGGACCTTCCTGCGCGTGGCCGCCCTGGCGCGCTA
1501     ---------+---------+---------+---------+---------+---------+  1560
         CGGCCGCCCGTGGGATGATCTCTCTCCCGACGAACAGGAGGCCCTTCGGCGCGAGGACGA
1561     ---------+---------+---------+---------+---------+---------+  1620
         AGACCGGTGGACCCGCCTCTCCAACCGCCAGAAGCGAGTGCGAAAGGCCAGTCAAAAACC
1621     ---------+---------+---------+---------+---------+---------+  1680
         TTTTCGGCTTTCCTTTGACGAGTGGCCAACTGAGGCTCGCAAAGAATGGGAGGACTACGA
1681     ---------+---------+---------+---------+---------+---------+  1740
         GCGCTATGCCTCATCGGCACCTGGGAGCATCGCGCGCGTGCAGGCGGCGCTTGCGGGCGC
1741     ---------+---------+---------+---------+---------+---------+  1800
         ACCTCTCGCTCCCACGACCGTGCGGACGGAAACGCTCGAGCGTGAGCGGATACTTATAGA
1801     ---------+---------+---------+---------+---------+---------+  1860
```
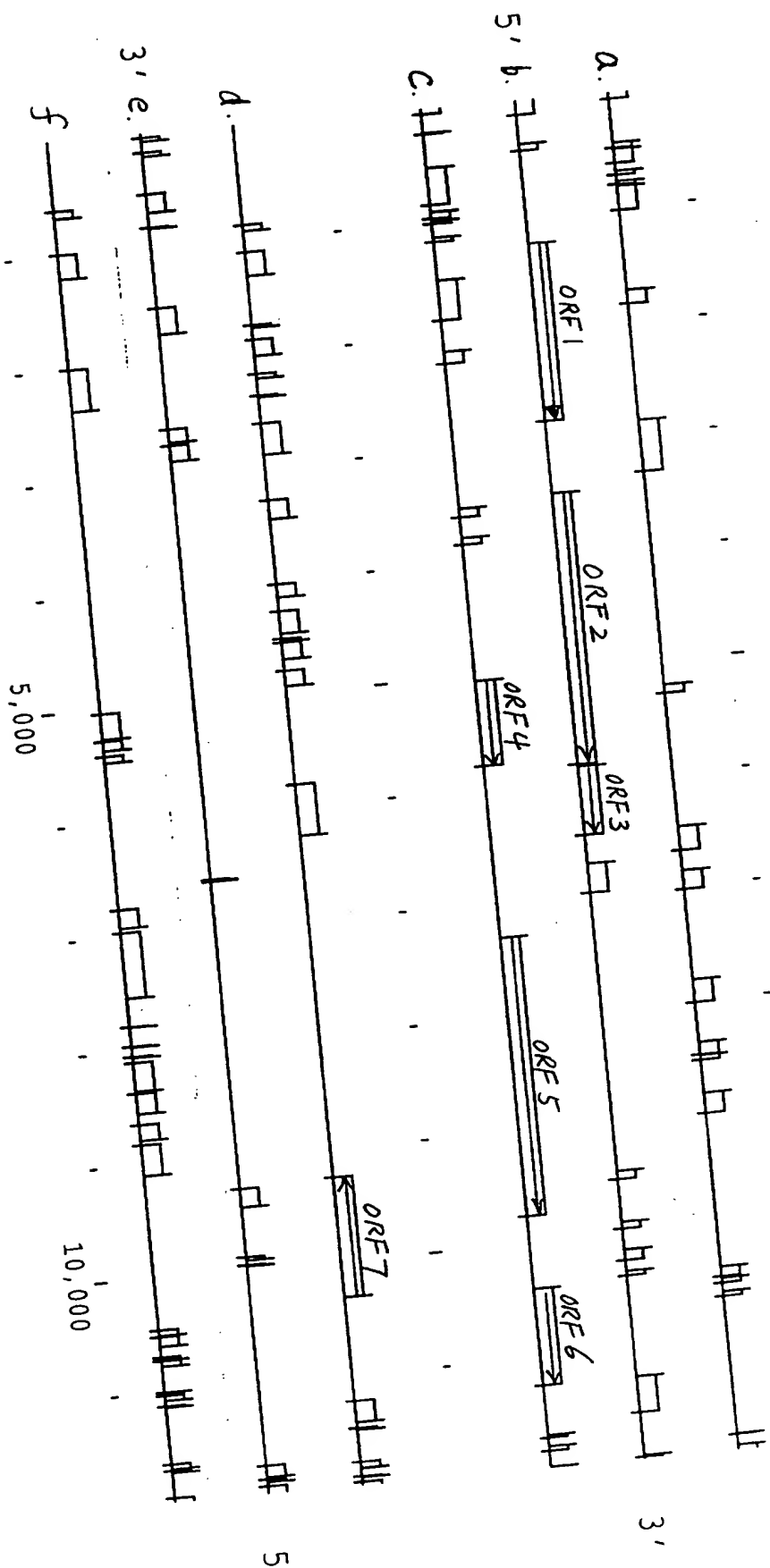
Fig. 7 (continued)

```
        ACTGTTCTACGGCTACTGTGTAAACGAACGGGGCCTCGACAGCAACGCGTTGAGCCTCGC
1861    ---------+---------+---------+---------+---------+---------+  1920
        CCTCCTCACAGACCTGGAGCTCGTCCAATCGTACCTGGAGTGGCGCGTGAATAGGTACAA
1921    ---------+---------+---------+---------+---------+---------+  1980
        GGACGAGGATTTACCCCCCGTTACTCGATCGGAATACATGTTTATCGCCCTGGTGAAAAA
1981    ---------+---------+---------+---------+---------+---------+  2040
        ACTCCACAGAGGTTATCTCCGCGCCCTTGGGCTTGGGGTAGACCCGGACGGGGTGAAAGA
2041    ---------+---------+---------+---------+---------+---------+  2100
        GCTGGAACGGAAACTGAAAATCGCCGGAATTGATGTCACGGACGGCTACCACGCGGTGGA
2101    ---------+---------+---------+---------+---------+---------+  2160
        GCCCCTCCTGGAAACTCACGAGCCCCTCCGCTGGGTGCTGGATGGCATCCGGCTCATGCT
2161    ---------+---------+---------+---------+---------+---------+  2220
        CCGCGATGCGGCGGGGCGGGTAGGCAACCTGCTGACACCCCAAATCCCCACCGCCAAAAG
2221    ---------+---------+---------+---------+---------+---------+  2280
        CGAAGCGGGCGAAGCGTTCGCCCTCTACCGGGACGTCGTTCTGCTTTGGATGATGGTGGG
2281    ---------+---------+---------+---------+---------+---------+  2340
        CCACCCCCTCCGGGCGAAGCATTACTACGAAGCTCGCTTGGACATGAGCCAGTTCCAAGA
2341    ---------+---------+---------+---------+---------+---------+  2400
        CGGGGATTTCGCTCCCGGGCGGGGACACGTGGGCGGGCCGGCGGAGGGTACTACCTGGC
2401    ---------+---------+---------+---------+---------+---------+  2460
        CTACCGCAAAGTGGAGTTCAAAAAACGCCCGAGGCCAGGTCTTTCAGAGCCTCCAGGACCA
2461    ---------+---------+---------+---------+---------+---------+  2520
        CGATCTCGTCACGTTCCCCCTGGACGACCCCGAGCACCCTGTCCTGGTCCTGGACGTGAA
2521    ---------+---------+---------+---------+---------+---------+  2580
        CGGGATGCGGTACTCCCTCAACGAGCTCTTTCACGTCTACCTGCGCACGATCCTCTCCCG
2581    ---------+---------+---------+---------+---------+---------+  2640
        CCTGGCCCAGGCCTGGGCCGGACCGGTCCCCTCCTGCCCCTGTTTCCGGGTGCCGATACG
2641    ---------+---------+---------+---------+---------+---------+  2700
        AGGCTCAGACTTGCGCACATCGTTCGCAGGCGCGCCGCCTACGTGGCCGCCGTGCCCGGG
2701    ---------+---------+---------+---------+---------+---------+  2760
        GTACCCCAGAAACTTTTGCCCTTCGGCCCCCACTCCATCCGCCACGTGGTGGCCACGGAG
2761    ---------+---------+---------+---------+---------+---------+  2820
        GTCGTGAAGCGCACGGGCTCTTTTGAGGCCGCCGCCAACGTGCTCCTGGATAGCATAGAC
2821    ---------+---------+---------+---------+---------+---------+  2880
        ATGGTCGTTCGACATTACGCCCGTTCGTTCCCCGCGACCGTAACAGTCACGGTTGGCGGG
2881    ---------+---------+---------+---------+---------+---------+  2940
        CTAACGCCCGCGCCCGGGGAGGTGAGCGGTGAGGGACCTCCACGACTTTTTCCTGGCCCG
2941    ---------+---------+---------+---------+---------+---------+  3000
        GGTGGACGAACTGGTGCCGGAACTCCTACCCGGGGCGCGGCGGGTGGGCGACGAGTGGCG
3001    ---------+---------+---------+---------+---------+---------+  3060
        GGCGGGCTCGGTCCAGGGCGAGCGGGGCGACAGCCTGGCCGTGGACCGCGGGAAGGGCTT
3061    ---------+---------+---------+---------+---------+---------+  3120
        CTGGATCGACCACAACCCCTCGGCCCCCCGAGCCCCGGCAGGGAAACCTCCTCACGCTGAT
3121    ---------+---------+---------+---------+---------+---------+  3180
        CCAGGCGGCCAAGGGGCTCTCCCCCGAGGAGGCCCGGCGCTGGGCCCAGCAGTGGCTTGG
3181    ---------+---------+---------+---------+---------+---------+  3240
        CCTCTCCCCTTCGCCAAAGGTCAGGCGGACGAGGAGCTCAGGACCAAAGGTCTTGAGTAC
3241    ---------+---------+---------+---------+---------+---------+  3300
        TCAAGTGCGTGGGAGCTCGGGTGCTCCAGTCCCTGAGTCTTCAGGTTCCCAGGTACCTGA
3301    ---------+---------+---------+---------+---------+---------+  3360
        GGAGTCGGACCCCTTTGACAACCCCCGCTTCCGGGACCTCCTCACCCCCAGGGGCGAGGA
3361    ---------+---------+---------+---------+---------+---------+  3420
        CGAGGCCCCCTTGGCCCCGGCCTCCGAGGAGGTGCTGCGGCGCATGGTGTCTAGGCTTCT
3421    ---------+---------+---------+---------+---------+---------+  3480
        CCGCACCCCCGAGGCCGTGGCCTACCTGAAGGGGCGCGGTCTGGATGCCCGGGTGGTCCG
3481    ---------+---------+---------+---------+---------+---------+  3540
        CCGCTTCTACCTCGGCCTGGACGACACCGCGCGGGCCACCGCCGCCCTGGTCTACCCGGT
3541    ---------+---------+---------+---------+---------+---------+  3600
        GATAGGGCCGGACGGCTCCCCCGTTCGCCGCCACCTCTACTACGAGATCCCCGGCCTCAC
3601    ---------+---------+---------+---------+---------+---------+  3660
        CCAGGGCGCCCCGGGCAAGGGCTGGGGGAGGGGGGAGGCCCACCAGCTACTGGGCCCTCCC
3661    ---------+---------+---------+---------+---------+---------+  3720
```

Fig. 7 (continued)

```
          CCCCTTCGAGGGCCCCTCCCCCCGCCGCAAGCTCTTCTTGTGCGAGGGGGCGAAGGATGC
     3721 ---------+---------+---------+---------+---------+---------+ 3780
          CTGGGCCCTCTGGCTCCACCTCCACGCCCAGCCCTGGGCCCAGGACCTGGCGGTGGTGAC
     3781 ---------+---------+---------+---------+---------+---------+ 3840
          CTCCACGCACGGCTCCGCCCTCCCCGAGGCCTGGAAAGACCCCCTGTTCTGGGCCCCTTG
     3841 ---------+---------+---------+---------+---------+---------+ 3900
          GGAGGAGGTCTACCTGGGCCAGGACGCCGACTCCGCCGGCGAGGAGATGGCCCGGAAGGT
     3901 ---------+---------+---------+---------+---------+---------+ 3960
          GGCGGAGGTGGCGAGGCGGCCCGTCCGCCGCGTCCGGGTCCCGGAGGGGATGGGGAAGGA
     3961 ---------+---------+---------+---------+---------+---------+ 4020
          CTGGACGGACTACTTCCTGGCGGGGGGCACCCCCGAGGGCTTGCGCCTCCTCCTGGAGGG
     4021 ---------+---------+---------+---------+---------+---------+ 4080
          AGCGGAGGTCTGGGAAGAAGAAGTGGCTGGAGGTGGGGCCAGGATCCAGCTCCCGGACCC
     4081 ---------+---------+---------+---------+---------+---------+ 4140
          CGTGGACATCCAGCGGGGCCTTCGTGCGGGGCCACCTCTACGTCCCCGTGCGGGTCCTGGA
     4141 ---------+---------+---------+---------+---------+---------+ 4200
          GAACCGGGGGGAAGAAGGGGCCCGCTACCGCACCGTGGTGGTCCGCTCCGACGGGGCCGT
     4201 ---------+---------+---------+---------+---------+---------+ 4260
          CCTGGGCTGGGGCTACTTGCCGGCCCCGCCCGGCACCCCCTTGGAGGACCGGGTGCTGGC
     4261 ---------+---------+---------+---------+---------+---------+ 4320
          CGTGGACGACGGCACCATCATCCGCAGGCCCCCGAAGGCGGCCGCCGGGACCTCGTGGAA
     4321 ---------+---------+---------+---------+---------+---------+ 4380
          CGGGGAGGCCATCAACCGCTTCCTGGAAGCCCGGGCCCGGGGAGTGAGCGCCATGACCGT
     4381 ---------+---------+---------+---------+---------+---------+ 4440
          GGCCCCCCGGGACCTGCCTGGGCTCATCGTCCGCCACCTCCGCCAGGTGATCCTCCCCAG
     4441 ---------+---------+---------+---------+---------+---------+ 4500
          TGAGGACGGCTACCTCCTGGCCGCCTTAGGGGTCATGACCTCCTACGTGCAGAGCGTCTT
     4501 ---------+---------+---------+---------+---------+---------+ 4560
          CGACGCCGTGCCCCTCTTCCTCGTGGTGGGCCCGCCGGGCTCGGGGAAGACGGAGTTCGC
     4561 ---------+---------+---------+---------+---------+---------+ 4620
          CCGCCTCATGGCCGAGCTGGGGGCCAACGGCGTGGTGATCACCGGCCAGACCTCCGCCGC
     4621 ---------+---------+---------+---------+---------+---------+ 4680
          CACCGCCGCCCGGATCATCGACGAGACGGGGGGGCTGGTGGCCTTCGACGACCTGGAGGA
     4681 ---------+---------+---------+---------+---------+---------+ 4740
          GGTGCGCCAGCGGTCGGGGAGCGCTGAGGCCTCCCAGCTGGAGCAGTTCCTCAAGGTGTC
     4741 ---------+---------+---------+---------+---------+---------+ 4800
          CTACAAGAAGGAGACCGCGGTCAAGAGCTGGACGGACACCAAGGGGATGCGGGTCCTCAC
     4801 ---------+---------+---------+---------+---------+---------+ 4860
          CCTCAACTTCTTCGGGGTCAAGGTGATCACCAACACCCAGGGGACGGGGGACATCCTGGG
     4861 ---------+---------+---------+---------+---------+---------+ 4920
          GAGCCGGATGCTGGTCATCCGCACCGCCCGCCTCCGGGACCTGGGCAGAGGGGAGGAGCG
     4921 ---------+---------+---------+---------+---------+---------+ 4980
          CCGCCCCGAGGGGCTCTCTCCCCCCCCAGGCCCTCCAAGAACTCCGGGACAACCTCTACATCT
     4981 ---------+---------+---------+---------+---------+---------+ 5040
          GGGCCATGGAGAACGCGGCCAGCCTCCACGCCCTGTACCGGGAGCGCTTCGCGGGCAAGG
     5041 ---------+---------+---------+---------+---------+---------+ 5100
          GGGAGCGCCTGGACGAGATCGCCGCCCCCTTGCGTACCATCGCCCACCACCTGGGGGACG
     5101 ---------+---------+---------+---------+---------+---------+ 5160
          AGGAGCTGGCGGCCCGCCTGGAGGACGCCCTGCGCCGGCAGGAAGGGCGCCTGGAGGAGA
     5161 ---------+---------+---------+---------+---------+---------+ 5220
          CCCTTTCCGATGCCGAGGTGGTGGAGACCGCCCTCAAGGAGGCCATCCGCCAGGGCTACC
     5221 ---------+---------+---------+---------+---------+---------+ 5280
          GGAGCCACGTGGCCCTGGTCCACGTGATCTTCCAGGCCCGGAAGATCTTCGGGGACGACT
     5281 ---------+---------+---------+---------+---------+---------+ 5340
          GGGGCCGGGAGCGCACCGTGGACATCCCCCGGTGGCGGGACCCCAAGTGGGTGGGGCAGA
     5341 ---------+---------+---------+---------+---------+---------+ 5400
          TCGCCAGCAACTACGGCTGGGCGGCCCCAGAAAGGCCCGTGAGGCCCCGGCTTTGGGACA
     5401 ---------+---------+---------+---------+---------+---------+ 5460
          AGCAGTTCCGCATCATGCGCCTGGAGCCCACCTTCGTGGAGCGGGTGGTCAGGGGCTTCC
     5461 ---------+---------+---------+---------+---------+---------+ 5520
          TCCAGGAGGGGATCCCCCTTGGAGCCCCTGAAGCAACCCCTGGCTTCTGCCTGGACACCCC
     5521 ---------+---------+---------+---------+---------+---------+ 5580
```

Fig. 7 (continued)

```
     CTGCGCCGAGTGCGCCTACCTGCACTGGTGCGACCTCCGGCCTGACAAGGAAAAGTGGCT
5581 ---------+---------+---------+---------+---------+---------+ 5640
     GGAGCGCTACGGGGAGGCCAAGCTGGCCCAGAAAAGGCGGGAGCTGGAGGAGGAGTTTTT
5641 ---------+---------+---------+---------+---------+---------+ 5700
     GGCCCTGGTGGGGCCCCAAGATGGCCTTGGCCTCCAGGCTTCCGCCGAGGAGGAGGGAGA
5701 ---------+---------+---------+---------+---------+---------+ 5760
     CCGAGGTAAGCACCCAAGTACCCAAGTACCCAAGACCCTAAAGCCTCAGGTACCGGAGGA
5761 ---------+---------+---------+---------+---------+---------+ 5820
     CCTCGGGGACGGAGGACCTAAAACCCCAAGGGCGTGAAAGACTGAGGTGAGAGGGATGAT
5821 ---------+---------+---------+---------+---------+---------+ 5880
     CGTGGCTGTCACCGGCTTCAAGGGAGGGGTGGGGAAGACCACCACGGCGGTCCACCTGGC
5881 ---------+---------+---------+---------+---------+---------+ 5940
     CTGCTTCCTGGCCGAGCGGGGCCCCACCCTGCTGGTGGACGGGGACCCCAACCGCTCCGC
5941 ---------+---------+---------+---------+---------+---------+ 6000
     CACGGGGTGGCACCGGAGGGGAGGCCTCCCGGTGACCGTGGTGGACGAGCGGGTGGCGGC
6001 ---------+---------+---------+---------+---------+---------+ 6060
     CCGGTACGCCCGGGAGCACGCCCACGTGGTCATAGACACCCAGGCCCGCCCCACGGAAGA
6061 ---------+---------+---------+---------+---------+---------+ 6120
     GGACCTCCGGGCCCTCGCCAAGGGGGTGGACCTGCTGGTCCTGCCCACGTCCCCCGACGC
6121 ---------+---------+---------+---------+---------+---------+ 6180
     CCTGGCCCTGGAGGCCCTCCTGGCCACCCTGGAAGCCCTGCGGGGGGGCGGAGGCCCGCTT
6181 ---------+---------+---------+---------+---------+---------+ 6240
     CCGGGTCCTCCTGACCATGGTGCCCCCGCCCCCGAGCCGGGACGGGGAGGAGGCCCGGGC
6241 ---------+---------+---------+---------+---------+---------+ 6300
     CCTCTTGGGGGCGGAGGGCGTTCCCCTCTTCACAGGCTGGGTGAGGCGGGCGGCAGCCTT
6301 ---------+---------+---------+---------+---------+---------+ 6360
     CCCCAAGGCCGCCCTCCTGGGGGTGCCCTGTCTACCGGGTGCCCGACCCCAGGGCGAGGCT
6361 ---------+---------+---------+---------+---------+---------+ 6420
     GGCCTGGGGGGACTACGCGCGGGTGGGGGAAGAGCTCCTGAAGGAGGTGGGGGGATGAGC
6421 ---------+---------+---------+---------+---------+---------+ 6480
     AAGTTCGCCAGGCTCCTCAAAGAGGTCAAGGAGAAGGAGGAGGCCTCCGGGGAGCGGCCT
6481 ---------+---------+---------+---------+---------+---------+ 6540
     CGGGGGAAGAGCCGGCGGGAGGACTACGTGGCCATGAAGGTCTACATCAGCAAAGAGCTT
6541 ---------+---------+---------+---------+---------+---------+ 6600
     CACCGGAGGCTGAAGCTGAAGGCCCTGGAGGAGGAGAAGGAGCTTTCGGAGCTGGTGGAA
6601 ---------+---------+---------+---------+---------+---------+ 6660
     GAGGCCCTGAGGAAGTTGCTGGTGTGACCTCCTCCCGCCTCGTAGAGCGTGAAAAGGAGG
6661 ---------+---------+---------+---------+---------+---------+ 6720
     TAAGACGATGGTCACCCTTAACAAATCGCCCCTAGAAGCCCTCTACGCGGGCCACTCCCC
6721 ---------+---------+---------+---------+---------+---------+ 6780
     CCAGGAGGCGGCCGTCTCTTCGAAGCGCCTGGTCCGCAAGATATTGAAGGAACTCCACC
6781 ---------+---------+---------+-----  ----+---------+---------+ 6840
     CCATCTGGAGCCAAGAGTTCGTGGATGTCGTCCCTTGGTCCGAGCACGCCACCGCAAGG
6841 ---------+---------+---------+---------+---------+---------+ 6900
     GGCTCAGGGCCACGGACATCGGCGTGGACCTGGTGGGCTACGGGAAGGACGACAAGGTCT
6901 ---------+---------+---------+---------+---------+---------+ 6960
     ACGCCATCCAGGTCAAGCTGTGGGATAAGCCCCTCTCTTGGAAGGACCTGGGGAGCTTCG
6961 ---------+---------+---------+---------+---------+---------+ 7020
     TGGGGGTGGTGAACCACCCCGAGTACGGCTTCGACCACGGGCTCATCGTGGCCCCAAGAG
7021 ---------+---------+---------+---------+---------+---------+ 7080
     GCGTGACCCAGGAGGCCGACCGCCAGCTCCAGGGCCTACCCATCACCATCCTGAGCGAAG
7081 ---------+---------+---------+---------+---------+---------+ 7140
     AGGCTCTCCTAGAAGACCTGGACCTGGAATCCCTCGTTCCAGACCGCCCCGAGGAAGCCC
7141 ---------+---------+---------+---------+---------+---------+ 7200
     GCAGGCGGGGGAAGAAGGCCCTCCGTAAGTACCAGCAAGAAGCCTTAGAGGAGGTGGCCA
7201 ---------+---------+---------+---------+---------+---------+ 7260
     AAGCCTTCTTAGAGAAGGGCCTGCCCCGGGGGCAAGCTCATCATGCCCCCGGGCACGGGCA
7261 ---------+---------+---------+---------+---------+---------+ 7320
     AGACCCTGGTGGCCCTCAAGATCGCCGAAAAGGTGGCGGGCCCCGGGGGGAGGGTCCTCT
7321 ---------+---------+---------+---------+---------+---------+ 7380
     TCCTGGCGCCCTCCATCGCCCTCCTGGACCAGTCCCTCAGGGCCTGGGCGGCGGAGGCTT
7381 ---------+---------+---------+---------+---------+---------+ 7440
```

Fig. 7 (continued)

```
       CCTTGCCCTTGCGCCTCTTCGCCGTGGTCTCGGACACGGGCGTGGGCAAGACCTCGGAGG
7441   ---------+---------+---------+---------+---------+---------+   7500
       ACGACCTCTCCGCCCTCTCCCTCCTCTCCATCCCTCCTACCACCAAGCCTGAGGAGCTGG
7501   ---------+---------+---------+---------+---------+---------+   7560
       CCTCCGAGGCCAAGACGGAGAGTCAGGAGGCCCTCACCGTGGTCTTCTCCACCTACCAGT
7561   ---------+---------+---------+---------+---------+---------+   7620
       CGGCGGAGGTCCTGGAGAGGGCCCAGAAGGAGCACGGGCTTCCCCCTTTTGACCTGATGA
7621   ---------+---------+---------+---------+---------+---------+   7680
       TCCTGGACGAAGCCCACCGCACAGCCACGGTGCGGGCGGGAGAAGAAAGCCCCTTCACCA
7681   ---------+---------+---------+---------+---------+---------+   7740
       AGGTGCACCACGACCACTACGTGAAGGCCCGCCACCGCCTCTACATGACGGCCACGCCCA
7741   ---------+---------+---------+---------+---------+---------+   7800
       GGATCTGGGAGGTGGAGGGGAATGGAGAGAGGGGCCAAGGGAAAAAGGCGGGGAAAAAGA
7801   ---------+---------+---------+---------+---------+---------+   7860
       AGGACCCTCAGAAAGAGGGTTCTCCTCCCCTTTTGGACCTCGGTGCCTCTCCTACGGAGG
7861   ---------+---------+---------+---------+---------+---------+   7920
       ACTCCACGGCCCCCGAAGGGGTGGAACTCCTGGTCTACTCCATGGACAACGAGGGGATCT
7921   ---------+---------+---------+---------+---------+---------+   7980
       ATGGCCCCACCCTCTACGAGTACACCTTCACCCGCGCCGTGAAGGAGGGCCACCTGAGCG
7981   ---------+---------+---------+---------+---------+---------+   8040
       ACTACAAGGTCATCGTCTTCTCCGTGGCGGAGGAAGCCCAAAAGGACCTGGCCTCCTACC
8041   ---------+---------+---------+---------+---------+---------+   8100
       TCCAGGGACCCGAGGCCCTCAAGGTGGAGGAGGCTCTGAAGGCCCTGGGCCTGTGGAAGG
8101   ---------+---------+---------+---------+---------+---------+   8160
       TCCTCCAGGGGGAGGTGCGGGACGAGGAGGGGAACCCGATGGGGGGGCCTCGACCTGCGGA
8161   ---------+---------+---------+---------+---------+---------+   8220
       GAGTCATCGCCTTCCACGGCCGGGTGAAGGAGTCCAAGGAGATGGAGGAAGAGTTCACGA
8221   ---------+---------+---------+---------+---------+---------+   8280
       AGGTGGCCCTCGCTGCCCAGCAGGCTGGCCTCCTTCCCGAGGAGCTCCGGCGGGTGGAGG
8281   ---------+---------+---------+---------+---------+---------+   8340
       TGAAGCACATAGACGGGCAGATGTCCGCCTATGACCGGAAGCGCCTCCTGGACTGGCTTA
8341   ---------+---------+---------+---------+---------+---------+   8400
       GGGAGAACGTCCCCGAGGGGGAGGTCCGCCTCCTCACCAACGCCAAGGTCCTCACCGAGG
8401   ---------+---------+---------+---------+---------+---------+   8460
       GGATCGACGTCCCCGGCCCTAGATGCCGTGGCCTTCATGCGTCCCCGGGACAGCGTGGTGG
8461   ---------+---------+---------+---------+---------+---------+   8520
       ACGTGATCCAGGCCGTGGGGCGGGCCATGCGCAAGGCCCCGGGCAAGGAGTACGGGTACG
8521   ---------+---------+---------+---------+---------+---------+   8580
       TGGTCCTGCCCGTGGTGGTGAGGGGGCAGGACGAGGAGCGGGAGATCGAGGAGAGCGGCT
8581   ---------+---------+---------+---------+---------+---------+   8640
       ACCGGGCGGTGTGGCAGGTGCTCTCGGCCTTGCGCTCGGTGGACAAGTCCTTCGAGGCCC
8641   ---------+---------+---------+---------+---------+---------+   8700
       GCATGCGCGCCCCCTGGTGCGCCTCTCGGGTAAGGGCGAGGGCGGGGAAGGTGGAGAGG
8701   ---------+---------+---------+---------+---------+---------+   8760
       CCCGAGAGGGTGTGGCCGTCATCGGGGAAGGAAGCGCCTCCCCCGTGATCGTAGATGTCC
8761   ---------+---------+---------+---------+---------+---------+   8820
       TTCAGGGGAACCTCAACCTCCACCAGGAGATCACCCGGAGCCTCGCCGGCAAGCTGGTCA
8821   ---------+---------+---------+---------+---------+---------+   8880
       GGCGCCTCGCCCTGGGGCGGAAGTACCTGGAGAACTGGGCCCAGGACGTGGCCCGGGTGG
8881   ---------+---------+---------+---------+---------+---------+   8940
       CGAAGGTGCTGGAGCAGCAGGTCAGGGCGATGGCGGAGCGGGACCCCAAGGTGAAGGAAA
8941   ---------+---------+---------+---------+---------+---------+   9000
       AACTGGGGAAACTCCTCGCCGCCCTGCAGGCCTTCACCAGCGAGAGCGTGACGGAGGACG
9001   ---------+---------+---------+---------+---------+---------+   9060
       AAGCCATCCTCATGCTGGTCCAGCACGCTCTCACCAAGCCCATCTTCGACGCCCTCTTCG
9061   ---------+---------+---------+---------+---------+---------+   9120
       GGGAACTCCTAGAAAAGCGGGAGGACCCCGTTTCCCGGGCCCTAGACGAACTCTTCCAGG
9121   ---------+---------+---------+---------+---------+---------+   9180
       AGTTCAGGGGGTTCCTGGACCGGGAAGGGGAGGCCCTCAAGGATTTCTACGAAGAGATGC
9181   ---------+---------+---------+---------+---------+---------+   9240
       GCCTCAAGGCCCTAGGGCTCACGGACGAAGCCGAAAGGGCCGACTTCCTACGGAGGCTCT
9241   ---------+---------+---------+---------+---------+---------+   9300
```

Fig. 7 (continued)

```
        ACTCCAACTTCTTCGCCCGGGCCTTCCCCCAGGTGGCCGACCAGGTGGGGATCGCCTACA
9301    ---------+---------+---------+---------+---------+---------+    9360
        CCCCGGTGGAGCTGGTGGACTTCCTGGTGAAGAGCGCAGACGAGCTGGCCAGGAAGCACT
9361    ---------+---------+---------+---------+---------+---------+    9420
        gTTGGCCGGGGGCTCGATGGGGAGAAGGTCTTCATCCTGGAGCCCTTCGCCGGCACAGGC
9421    ---------+---------+---------+---------+---------+---------+    9480
        ACCTTCGTCACCCGAATCCTGCACCGGGTAGCCGAAAgGGGCGGGGCCGACGCGGTCAAG
9481    ---------+---------+---------+---------+---------+---------+    9540
        GGCAAGCTGGAGCGGGGGGAGATCTGGGCCAACGAGATCCTTCTCCTCCCCTACTACGTC
9541    ---------+---------+---------+---------+---------+---------+    9600
        CTCAGGGCCAACGTGGAGAACACCACCCTGGCCCTGACCGGGGAGTACGTCCCCTTCAAG
9601    ---------+---------+---------+---------+---------+---------+    9660
        GGGGCGTTCTGGCGGACTCCTTCGGCTGGCGGAGCTGGGGTATAGCGAGAAAAAGTTTGG
9661    ---------+---------+---------+---------+---------+---------+    9720
        CATCATCCCGCTCTTCCCGGAAGAATACGGTGAGGCCCTGAACGAGCAGCTGAAGGCCCC
9721    ---------+---------+---------+---------+---------+---------+    9780
        TATCCAGGTTATCCTCTCCAACCCCCCGTGCGGGCTTGGTTGGAGAAGGAGGGCGAGGGG
9781    ---------+---------+---------+---------+---------+---------+    9840
        AAGAAGAACCCCGTCTACCGTAAGGTGCGGGAGCGGGTGGAGCCAACCTATGTACGGCGG
9841    ---------+---------+---------+---------+---------+---------+    9900
        GCCAAGGAACTTCCCATCGGGGGGACAAAACCCAAGGGAGAGAACCTGAACTCCCTCTAC
9901    ---------+---------+---------+---------+---------+---------+    9960
        GACCAGTACATCCAGGCCTTGCGGGTGGCGAGCGACCGTATCGGGGAGGAGGGGGTCGTG
9961    ---------+---------+---------+---------+---------+---------+    10020
        GCCTTCGTCACCCAACAACGGGTGGCTGGGGGGCGTAGTGCCCCGGGGCTTGCGGGCCTCT
10021   ---------+---------+---------+---------+---------+---------+    10080
        TTGGCGGAGGAGTTCGCCGAGGTGTACGTCTACGACCTGAGGGGGGGATGCGAGGGAGAAG
10081   ---------+---------+---------+---------+---------+---------+    10140
        GGGGAGGCACGGAAGAAGGAGGGGGGCGGGGTCTTTGGACAGCCTTCCCGCGCCGGGGTC
10141   ---------+---------+---------+---------+---------+---------+    10200
        TGCCTCCTCCTCCTGGTGAAGCGTAAGGACCACAAAGGGATCGGCAAGGTCCACCTCTAT
10201   ---------+---------+---------+---------+---------+---------+    10260
        CGGGTCGGGGACGGCCTCTCCCGGGAGGCCAAGCTGGCTCTGGTGAAGGAGCATGGCTCA
10261   ---------+---------+---------+---------+---------+---------+    10320
        GTCTCTGGGTTCCCTGGCAAGAGGTTCCCTATGAAGAGTGGGTGGGGAGGCTTACCCCCG
10321   ---------+---------+---------+---------+---------+---------+    10380
        GGTTCTCGGGGATGTTGTCCCTGGACGAGGTCTTTGAGGTGCGGAGTTCTGGGGTGAAGA
10381   ---------+---------+---------+---------+---------+---------+    10440
        CCAACCGCGATGCCTACGTCTTCAACCCCTCCCGGGCGGAGCTGGAGCGGCACATGAGGC
10441   ---------+---------+---------+---------+---------+---------+    10500
        GGCTCATCTCCACCTACAACGAGCACGTGAAAAGGAAAAAAGAGGGGAAACTAGGGGAAC
10501   ---------+---------+---------+--------- ----:---------+---------+    10560
        TGGAAAAGGATGAGAGCATCATCAAGTGGGATAGGGAACTCATCAGGTACCTAGAGTCCC
10561   ---------+---------+---------+---------+---------+---------+    10620
        TGAGGGAAGCTTCCTACGAAGGGAGCGGTCAAGTCTACGAGGCCCTCTACCGCCCCTTCG
10621   ---------+---------+---------+---------+---------+---------+    10680
        TGCCTATGTACCTCTACCTCAGCCGCACTTTCAATAGCATGATTTACCAAATCCCCCGCA
10681   ---------+---------+---------+---------+---------+---------+    10740
        TCTGGCCCACCCCCGAGGCCGAGAACCTGGCCATCGCCGTGGCCGGAAAGGGGAGTAACG
10741   ---------+---------+---------+---------+---------+---------+    10800
        CTTTTAGCGCTGTGGCCACCAGGAGGGTGGTTGACCTGCACTTTATTGAGACCACCCAGC
10801   ---------+---------+---------+---------+---------+---------+    10860
        TCTACCCCCTTTACCACTACCCCGAAAACAGCCCTCTGGGGGGACACCCAAAGCGCAAGC
10861   ---------+---------+---------+---------+---------+---------+    10920
        TCAACCTCAAGGAGGAGTTCTTGAGGAAGCTTGGGGAGGTCCTCGGCCGCCCCGTTCCCC
10921   ---------+---------+---------+---------+---------+---------+    10980
        CCGAGGAGGCCTTCGCTTACATCTACGCCGTGGTGAGCCACCCCCTCTACGCCGAGCGCT
10981   ---------+---------+---------+---------+---------+---------+    11040
        TCGCCAAGGACCTCAAGATGGACCTCCCCCGCATTCCCCTCCCCCAAGATCCCGAACTCT
11041   ---------+---------+---------+---------+---------+---------+    11100
        TTGCCAGGCTGGTGAAGGCGGGTCAAGAACTCATTCACCTCCACACCGAGTACGAGACCC
11101   ---------+---------+---------+---------+---------+---------+    11160
```

Fig. 7 (continued)

```
       TGCCCCCCTGGAGCCCAGTCCCCCTTCGGGTGGAAGAGGGAGGCCCGGAGGACCCTACGA
11161  ---------+---------+---------+---------+---------+---------+  11220
       GCGCTACCGGGTGGAGCGGATGAGGCTGGACAAGGAGAGGAGGGTTCTCCAGTACAACGA
11221  ---------+---------+---------+---------+---------+---------+  11280
       CTGGGTCCGGGTGGAGGGCATCCCCGAGGAGGCCTTCCGCTGGCGCCCCGGGGGGTACTC
11281  ---------+---------+---------+---------+---------+---------+  11340
       CCCCTTGGAGTGGATTGGCCGCTTCTGGAAGGTGGAGGAGAAGGTGCCCAAGGGCAGGGG
11341  ---------+---------+---------+---------+---------+---------+  11400
       GGAGGCCATCGTCTGGGACCCCAACCTCTTCCTCAAGGAGAAGGGGGAACCCCGTTACCT
11401  ---------+---------+---------+---------+---------+---------+  11460
       CCTGGACCTCATCGGGCGGGCGGTCCAGGTGGCCGTGCAGACGGTTGGGATCCACGAGGA
11461  ---------+---------+---------+---------+---------+---------+  11520
       GCTGAGAGAAGACGTGGAAGCTCTGCTGGGTTGAGGGGGTGCTGGCCCGCCGTTCTCCCT
11521  ---------+---------+---------+---------+---------+---------+  11580
       ACTCCTTTAGGGCCTACCCCTACGATCCAAGCACGGCCCTGGGGGGCGCTCAGGTGGGCA
11581  ---------+---------+---------+---------+---------+---------+  11640
       TCCCACGTCCAAGGCCCCGACTTGGGCACCCCATGCTGCGAACTTACAGCCCAAGGGCCT
11641  ---------+---------+---------+---------+---------+---------+  11700
       GAAACATTCCCCCCTGCTCACGGGGGAAAGTTCGTGAAGGAAAGAGCAAAGCCTTTTTTA
11701  ---------+---------+---------+---------+---------+---------+  11760
       TCGCATCCGGAGAGATGGCGGGGTGGAACTTTTCCCCGAGGACTCCCCCATAGGGACATG
11761  ---------+---------+---------+---------+---------+---------+  11820
       TAAACGGCAAGCTATCAGTGTAGACTTTTTTCAAAAAGAGCCATACTCGTGTTTTCCCGT
11821  ---------+---------+---------+---------+---------+---------+  11880
       TCAGAACGGCATTTTTGCTAAGGAGGTGGTTTACAAATGGGTGTTAATGCGCTACATCCT
11881  ---------+---------+---------+---------+---------+---------+  11940
       CCGGTAGTAGGAGCATGC
11941  ---------+--------  11958
```